# Integrating the New York Citywide Immunization Registry and the Childhood Blood Lead Registry

Hadi Makki, NYC DOHMH

hmakki@health.nyc.gov

Vikki Papadouka ,Ph.D, NYC DOHMH

vpapadou@health.nyc.gov

Paul Schaeffer, MPA, NYC DOHMH

pschaeff@health.nyc.gov

Amy Metoka, NYC DOHMH

ametroka@health.nyc.gov

# Objectives

- To describe the NYC Department of Health and Mental Hygiene's (DOHMH) Master Client Index (MCI) centralized de-duplication service

- To describe the integration of the first two systems to join the MCI, the Citywide Immunization Registry (CIR) and the Blood Lead Registry (LQ)
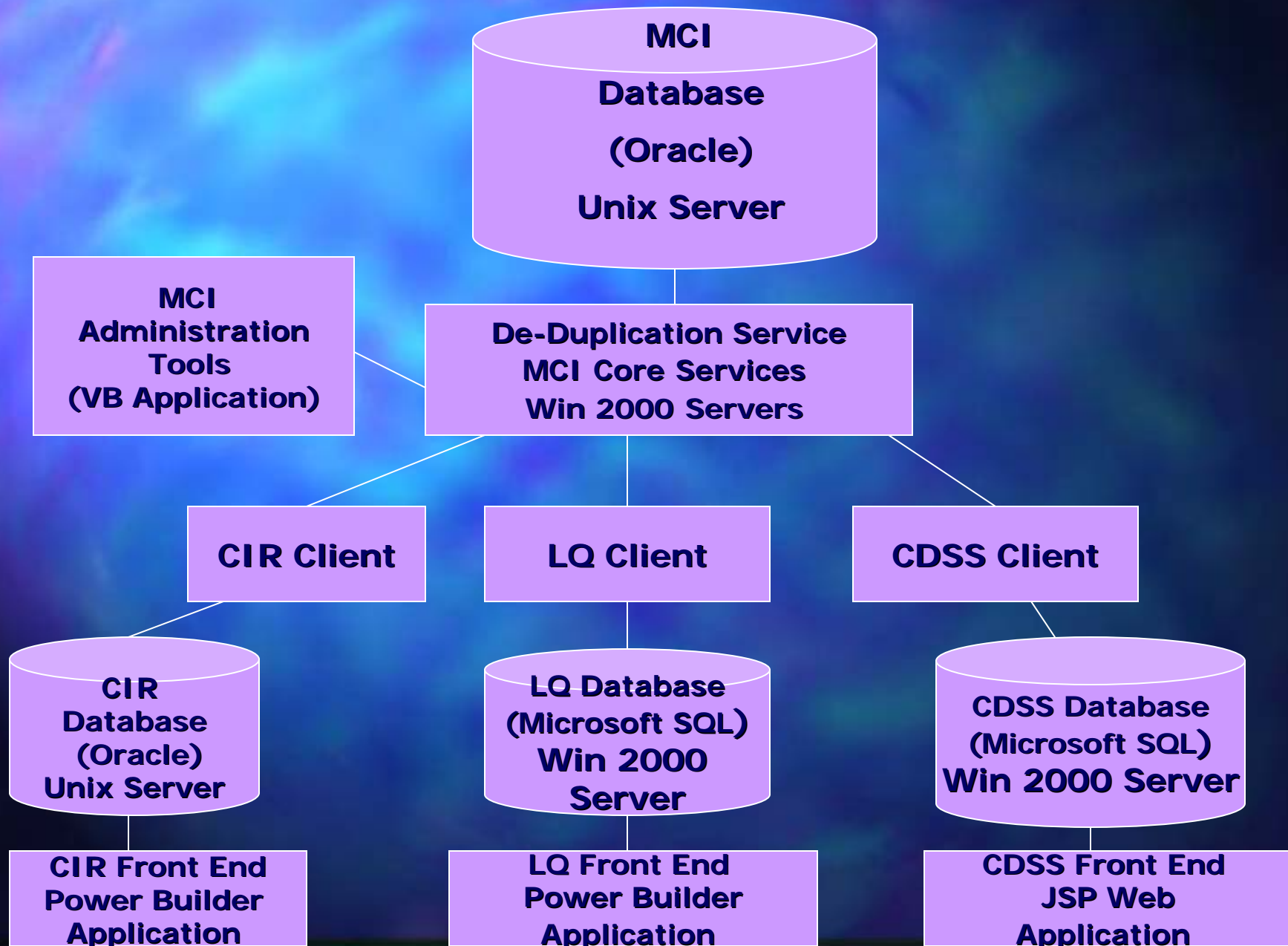
# Background – MCI

- The MCI integrates data from and provides a centralized de-duplication service to:

    - Citywide Immunization Registry (CIR)
    - Lead Quest Registry (LQ) from the Lead Poisoning and Prevention Program
    - Communicable Disease (2004)
    - Additional databases (in the future)

# What is the MCI?

- MCI – database that stores information from different programs for matching

- Core Services
  - -Business Rules
  - -De-Duplication Services

# Master Client Index

**MCI**

**Database**

**(Oracle)**

**Unix Server**

**MCI Administration Tools (VB Application)**

**De-Duplication Service**
**MCI Core Services**
**Win 2000 Servers**

**CIR Client**

**LQ Client**

**CDSS Client**

**CIR Database (Oracle) Unix Server**

**LQ Database (Microsoft SQL) Win 2000 Server**

**CDSS Database (Microsoft SQL) Win 2000 Server**

**CIR Front End Power Builder Application**

**LQ Front End Power Builder Application**

**CDSS Front End JSP Web Application**

# De-Duplication : Features and Process

- Potential duplicate pairs are reviewed by humans to train the model and "AI" model created, which contains "clues"

- Match thresholds are determined

- Incoming Records to MCI (through clients)

- De-Duplication happens in MCI and propagates to client systems

# CIR and LQ Integration

- In February 2004 DOHMH completed the integration of the first two systems to join the MCI

- The CIR and LQ each contains over 2 million children

# Results of Integration

**CIR**

| | |
|---|---|
| N pre-MCI | 2426369 |
| N post-MCI | 2065230 |
| N merged | 361139 |
| % merged | **14.9%** |

**LQ**

| | |
|---|---|
| N pre-MCI | 2184216 |
| N post-MCI | 2021635 |
| N merged | 162581 |
| % merged | **7.4%** |

# Results of Integration

| Birth Cohort | CIR | LQ | Merges | % LQ with CIR |
|---|---|---|---|---|
| <96 (No VRs) | 851460 | 1235734 | 494595 | 40.0% |
| 96-03 | 1210956 | 785685 | 636830 | 81.1% |

Note: CIR & LQ have different populations for children whose DOB < 1996

# De-Duplication – Specifics

- Within system automatic merging was 14.9% for the CIR and 7.4% for LQ

- Over 80% of LQ records merged with CIR records in children born between 1996 and 2003 (CIR has VRs)

- CIR duplication rate will be reduced from 30% to under 12% (with human review)

# Specific Benefits – CIR and LQ

- Improved Data Quality:  Complete Records – no longer fragmented

- 12 CIR and LQ FTEs currently needed for Human Review will be reduced by at least 50%

- CIR now disseminates Lead Test information to MCOs and providers on the phone and online - potential for greater public health impact

# Benefits – Centralized De-Duplication Service

- Cross-program leveraging of resources

- Programs can be granted access to other program's data

- Less FTEs needed for human review – re-deployment of staff

# Future Plans

- Integrating the MCI into the NYC DOHMH infrastructure as a central web-based de-duplication service